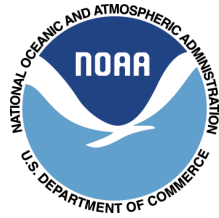# NGGPS Post-processing

Overall Objective

– Improve post-processing methods, data sets and products, resulting in products that are dramatically more useful to the weather enterprise

Strategy

– Query experts and stakeholders on what's needed

– Evolve supporting infrastructure to better support research to operations (R2O)

– Identify and generate supporting data sets

– Identify and develop advanced post-processing techniques

NGGPS Post-processing

Focus Areas

- Evolution to a community, open-source post-processing infrastructure in response to stakeholders.

- Generation of global reanalyses and a durable infrastructure that enables regeneration when needed.

- Global reforecasts based on reanalyses.

- Development of advanced post-processing techniques that support the National Blend of Models (NBM)

- Development of post-processing techniques for weeks 2-4

Input: UMAC, August 2015

UMAC | UCACN Model Advisory Committee

- UCAR Community Advisory Committee for NCEP (UCACN) established in March 2011 as a permanent external advisory committee to provide guidance to NCEP.

- UCACN Modeling Advisory Committee (UMAC) provides a comprehensive, technical review of the NCEP Production Suite (NPS) strategy for development

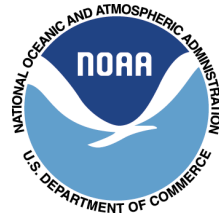- UMAC recommendations addressed post-processing

UMAC recommendations

- General: evidence-based decision making, reduced complexity of production suite, understand requirements, collaboration beyond NWS, actionable strategic and implementation plans.

- Reanalysis, reforecast (R/R), post-processing specific recommendations.
  - Regular production of R/R recognized as integral to production suite, so plan HPC/disk resources for doing this regularly in the future. Use non-operational HPC for their generation (cheaper).
  - However, extent of R/R and when to do a new R/R should be based on evidence, not political pressures.
  - Migrate post-processing resources to National Blend.
  - Also: develop techniques that require less training data; engage international partners to expand MME; hire ~5 Ph.D.-level statisticians to bring statistical rigor. Consider co-locating MDL and NCEP. Invest in development of hi-res. analyses for training and verification.

Input: Workshop, January 2016

Workshop on statistical post-processing, held in College Park MD, Jan 2016.

Included > 90 participants from NOAA, labs, universities, other met services, and corporations.

Participants expressed strong support for community-based efforts for StatPP across NOAA, and the wider weather-climate enterprise.

Post-processing Challenges

Note the substantial overlap.

Lack of community infrastructure

Organization

Science

Getting the right data in place.
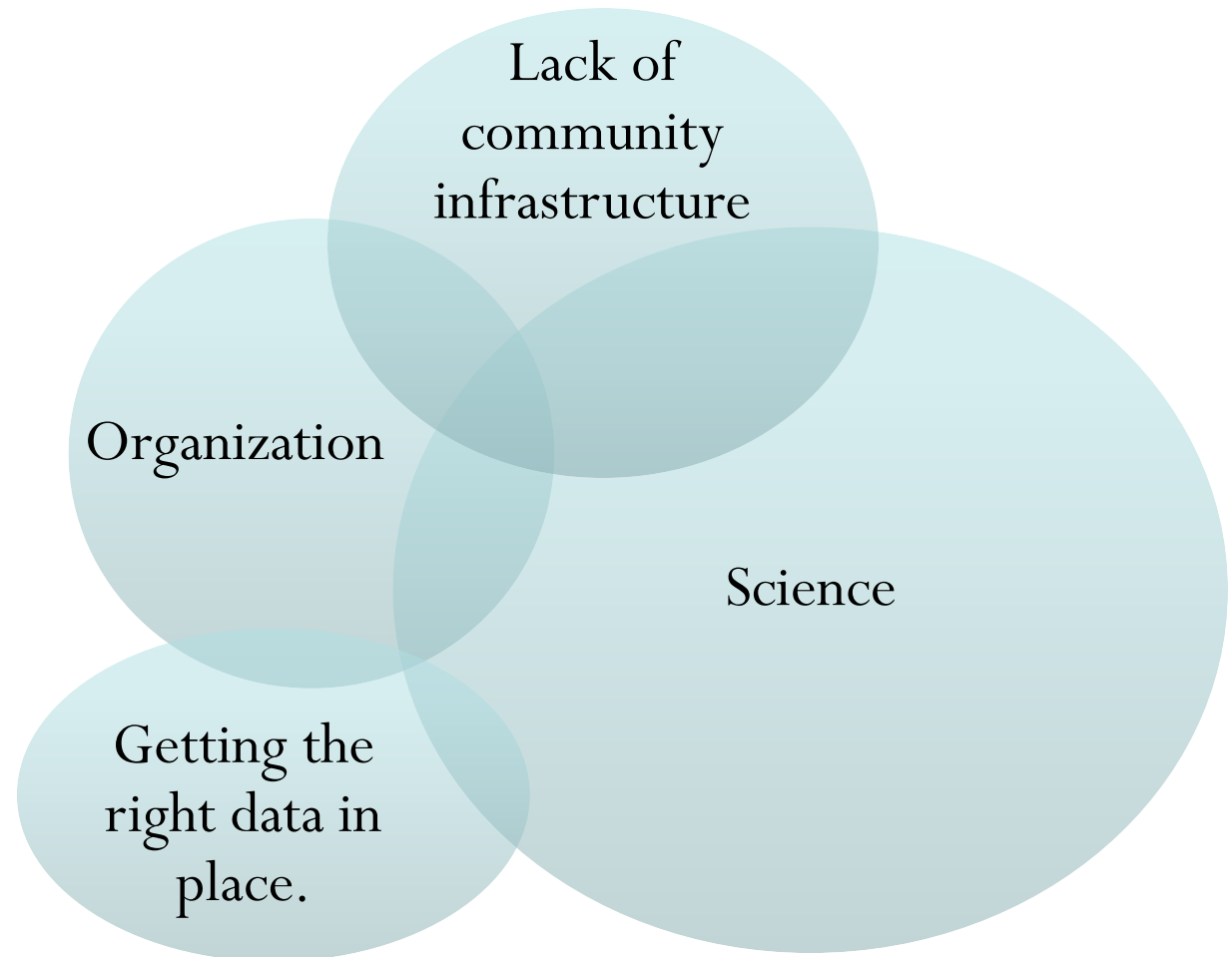
Post-processing workshop recommendations (1)

- Full recommendations [here](#).

- **Science**:
  - >4 professional statisticians.
  - Determine best algorithms for "foundational" guidance (e.g. T2m, POP, etc.) and use them.
  - Generate R/R's regularly.
  - Generate high-quality, hi-res analyses for training/verification.
  - Determine how to efficiently use the least amount of training data.

- **Organization**:
  - Define requirements, per UMAC.
  - Develop and use strat/implementation plan.
  - Develop a new, open-source, community PP infrastructure.
  - Train staff in software, modern languages.
  - Organizational streamlining.

Post-processing workshop recommendations (2)
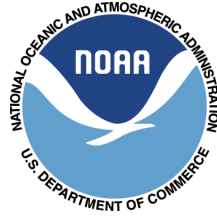
- **Community infrastructure**: Establish modern, widely-utilized distributed version control system to facilitate R2O and O2R.
  - Use industry-standard version control software, e.g., git.
  - Develop a code repository with tiers (outer: all can contribute; inner: tightly controlled by NOAA).
  - Governance process for community participation, migration of code from outer to inner.
  - Metadata, test data, and standards established.
  - More (see full recommendations)
- **Data and infrastructure**:
  - R/R incorporated into HPC planning.
  - Make foundational PP guidance readily available to wider community.
  - Global R/R data sets, analyses, etc. also made readily available.

NGGPS Post-processing Summary

- **Major Accomplishment in FY16:**
  - Held successful post-processing workshop in January 2016. Briefed outcomes to NWS leadership and wrote Functional Requirements Document for Statistical Post-processing (StatPP)
  - Developed a "Functional Requirements Document for Post-processing" and delivered to STI management.
  - Next-gen R/R project for GEFS underway.
  - Rollout of basic algorithms for National Blend.
- **Priority Focus for FY17**
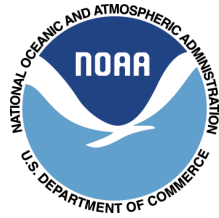 - Generate global reanalysis/reforecast and set up durable infrastructure to regenerate future reanalyses.
  - Develop Weather Information Statistical Post-processing System (WISPS), the community-based infrastructure.
- **Key Issues**
  - Management attention to the UMAC and PP workshop recommendations.
  - Develop software solutions that facilitate R2O and O2R

NGGPS Projects

Toward the Operational Production of Next-Generation Global Reanalyses

Tom Hamill (ESRL/PSD), with contributions from NCEP/EMC and NCEP/CPC

- This project will generate a multi-decadal reanalysis and reforecast for the GEFS system. Funding also via OAR/CPO.

- Year 1 (current) of the project is setting up the infrastructure and testing it at lower resolution during periods of changes in the observing system, and testing methods for breaking production into streams.

- Year 2 will largely be dedicated to the production of reanalyses and then reforecasts.

- Main challenge has been getting our staff access to the WCOSS system; > 6-month delay for security clearances.

- Reforecasts are needed for high-quality post-processing of rare events (e.g., heavy precip) and longer-lead forecasts (e.g, weeks 2-4 T, P). Reanalyses are needed for their initialization.
- R/R generation is an important part of the NGGPS post-processing implementation plan.

National Blend of Models Probability of Precipitation and Precipitation Amount
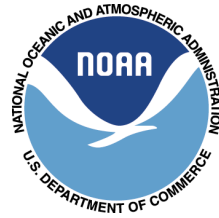
Tom Hamill PI, ESRL/PSD

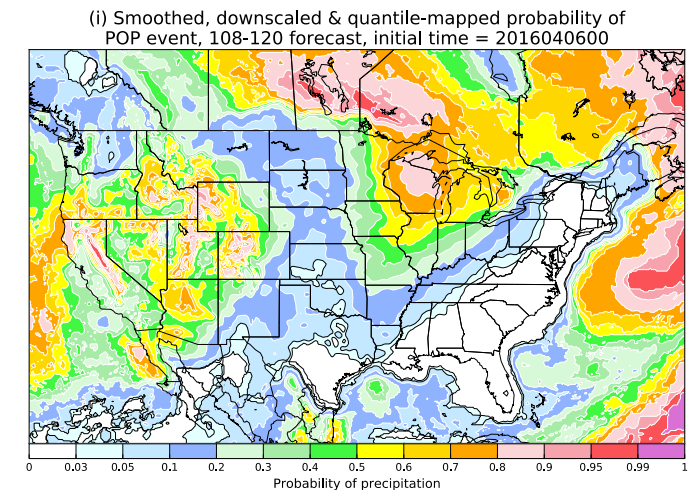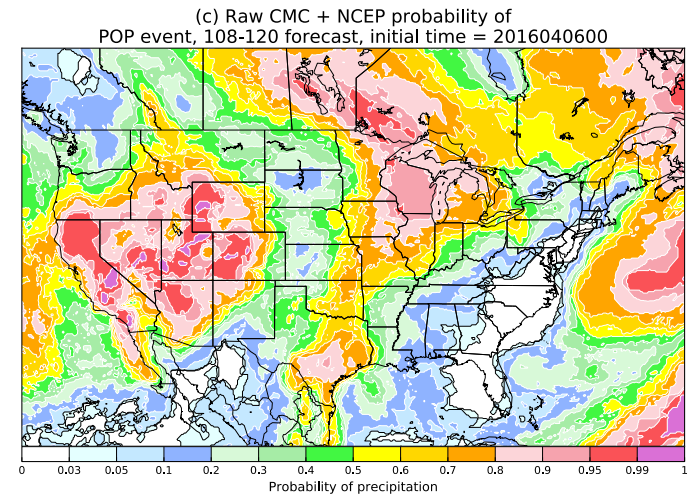- Via funding from many sources, NGGPS included, ESRL/PSD is delivering an advanced 12-hourly POP and 6-hourly QPF algorithm to MDL for use in the National Blend of Models.

- Implementation expected for v. 2.1 in Dec 2016

- The POP's are quite reliable and sharp, the QPFs are designed to provide useful guidance on precipitation maxima.

- The NGGPS post-processing implementation plan has as one of its major objectives the development of advanced post-processing methodologies, especially for high-impact variables, and their implementation in the National Blend of Models program.

(c) Raw CMC + NCEP probability of POP event, 108-120 forecast, initial time = 2016040600

0   0.03   0.05   0.1   0.2   0.3   0.4   0.5   0.6   0.7   0.8   0.9   0.95   0.99   1
Probability of precipitation

(i) Smoothed, downscaled & quantile-mapped probability of POP event, 108-120 forecast, initial time = 2016040600

0   0.03   0.05   0.1   0.2   0.3   0.4   0.5   0.6   0.7   0.8   0.9   0.95   0.99   1
Probability of precipitation

Tom Hamill, ESRL/PSD

- At the heart of POP12 is "*stochastic quantile mapping*." This both corrects for conditional bias (the bias for a given forecast precipitation amount) and adds noise from a normal distribution whose standard deviation depends on the mapped precipitation amount. This procedure is applied to each ensemble member. POPs are calculated from the ensemble and smoothed.

Stochastic quantile mapping for 6 Apr 2016: 120-132 h forecast, lon = -123.75, lat = 47.75

- Pending funding, ESRL/PSD will continue to evaluate alternative POP12 methodologies, including Michael Scheuerer's CSGD (censored, shifted Gamma distributions), and work with MDL to transfer the best possible algorithms for operational use.
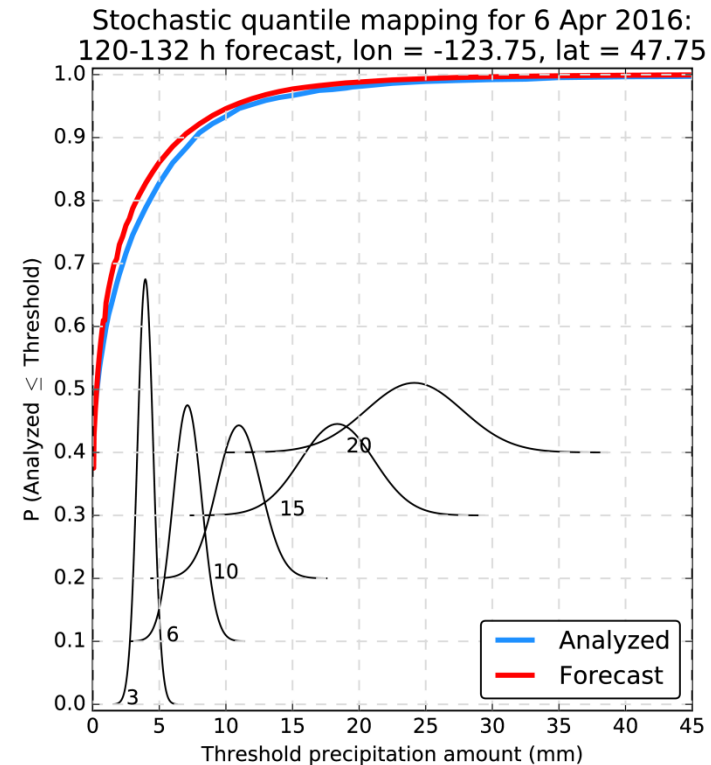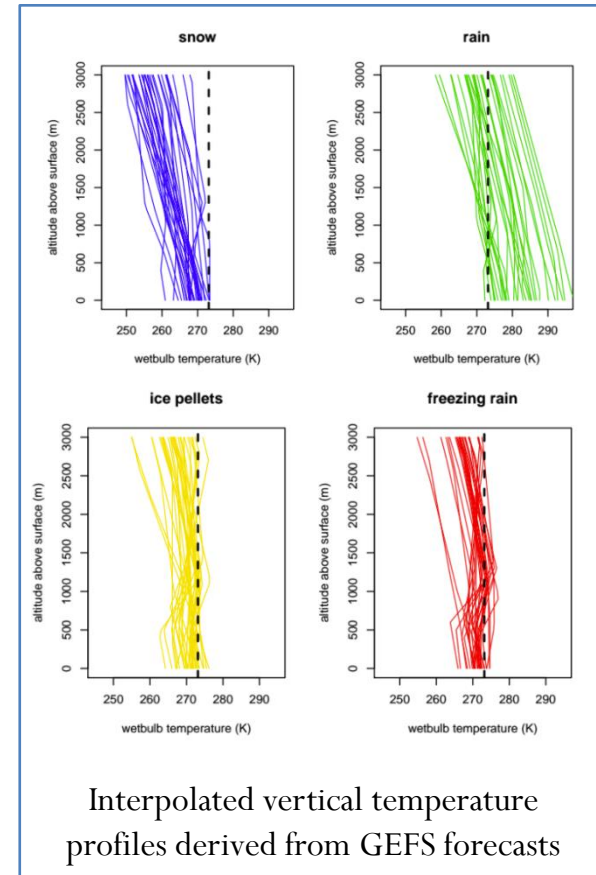
13

Probabilistic Forecasts of Precipitation Type and Snowfall Amounts based on Global Ensemble Forecasts
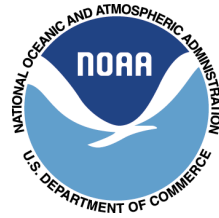
Thomas M. Hamill, ESRL/PSD

- Predictors: GEFS model predictions of surface and pressure level temperatures

- Objective: reliable probabilistic predictions of precipitation type and snowfall amounts

- Collaboration with MDL staff permits comparison against existing methods and may inform future operational use

Interpolated vertical temperature profiles derived from GEFS forecasts

- The NGGPS implementation plan promotes the development of improved post-processing algorithms for high-impact events
- Freezing rain and heavy snowfall can have a high impact, but have not been studied as much as other variables
- This project explores new statistical techniques, focusing especially on improving forecasts of freezing categories

Probabilistic Forecasts of Precipitation Type and Snowfall Amounts based on Global Ensemble Forecasts

Thomas M. Hamill, ESRL/PSD

- The new precipitation type algorithm yields reliable probability forecasts for all types

- Its performs similar to MOS for snow/rain, for freezing types it performs better

- The work on the algorithm for snowfall amount is ongoing

- A recent decision of the NBMWxG Team requires forecasts to be based on a number of specific input variables that can be edited by WFO forecasters
- Since the method studied here uses the full vertical profile, it does not meet these requirements
- Our developments could, however, inform future operational implementations, if the requirement of forecaster editable grids becomes obsolete one day
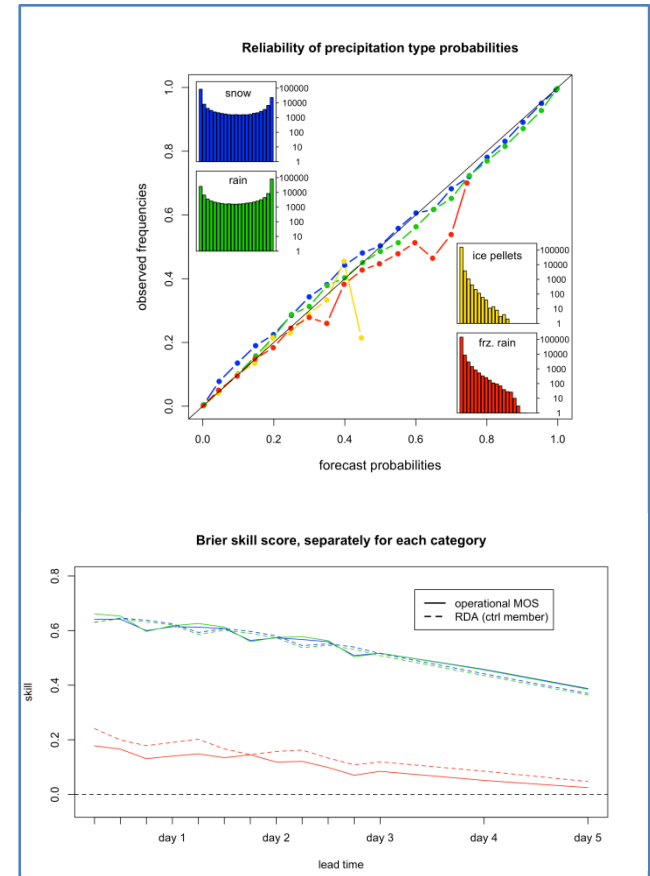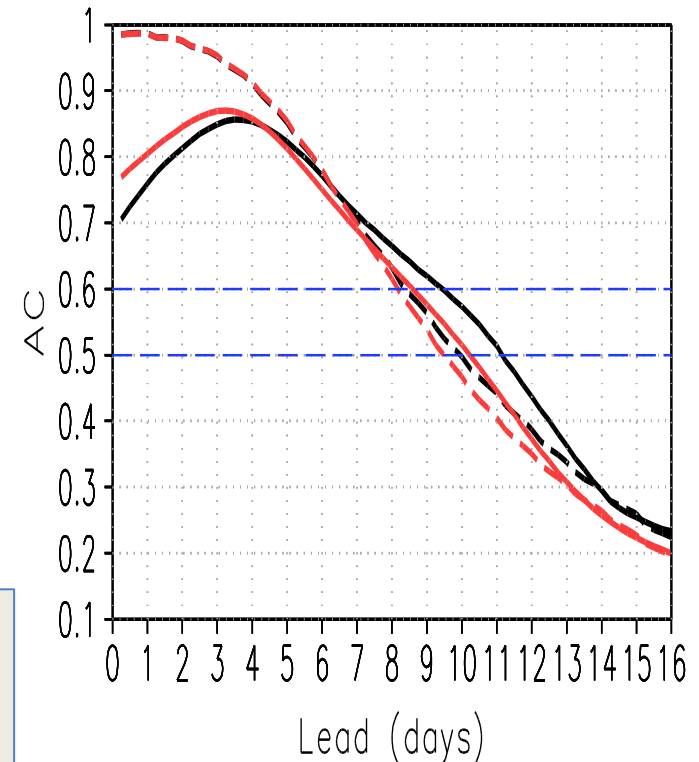
Ping Liu, Stony Brook University

- Separating lasting signals from chaotic noises with an EOF approach proves effective to improve the week-2-4 forecasts. Persistent ridges and blockings (PMZs) and MJOs are the two signals investigated in this project.

- A new tracking algorithm is for monitoring PMZs; a revised RMM index is for MJOs.

- Connection to NGGPS

  The tracking algorithm of PMZ, the revised RMM index of MJO, and the EOF approach being investigated on the GEFS ensemble forecasts will be applied to the NGGPS to improve the predictions at weeks 2-4.

Useful skill of Z500 associated with the PMZs can be extended by 1.2 days (solid vs. dashed in black) in the Northern Hemisphere after the EOF approach.
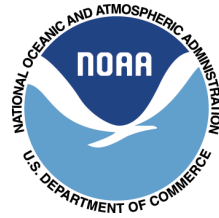
Ping Liu, Stony Brook University

- Summary research project results

A Revised RMM index (*Mon.Wea. Rev.*, 2016) is developed to observe the MJO in real time and in GEFS ensemble predictions.

A new tracking algorithm of PMZs (submitted to *Mon.Wea. Rev.*) is observing persistent ridges and blockings in real time and in GEFS ensemble predictions. The algorithm has identified substantially more persistent ridges to impact the T_{2m} and precipitation over the US during all seasons.

- Potential transition to operations/future implementation

The Revised RMM index can compensate or replace the existing RMM index in evaluating the ensemble predictions of the GEFS and NGGPS.

The new tracking algorithm of PMZs can compensate or replace the existing indices for identifying the blocking episodes in the predictions of the GEFS and NGGPS.

These two indices are updated daily on a web server constructed in collaboration with the EMC and CPC. The web pages are nearly ready in transition to both EMC and CPC.

Improved Statistical Post-processing with the **Bayesian Processor of Ensemble (BPE)**
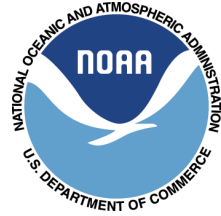
Zoltan Toth (GSD/ESRL), Mark Antolik (MDL/NWS), Roman Krzysztofowicz (Univ. Va), Melissa Petty (CIRA at GSD), Geary Layne (CIRES at GSD), and Malaquias Pena (IMSG at EMC)

- Fuse climatology & multiple forecast guidance
  - Prototype for probabilistic National Blend of Models
- Bayesian & distributional approach
  - Extremes handled well with smaller set of hindcasts
  - All continuous variables calibrated with unified method
- Multiple & versatile output formats
  - Probabilistic NDFD w only factor of 2 storage increase
- Algorithms & codes to be developed & tested
  - Comparison with operational EKDMOS

- Contributing to accomplishing major NGGPS goals
  - "Reliable weather, water… information"
  - "Improv[e] the probabilistic guidance of high impact weather elements by
    - Decreasing… uncertainty and improving product calibration"
  - "Cutting edge techniques" made available to NWS to
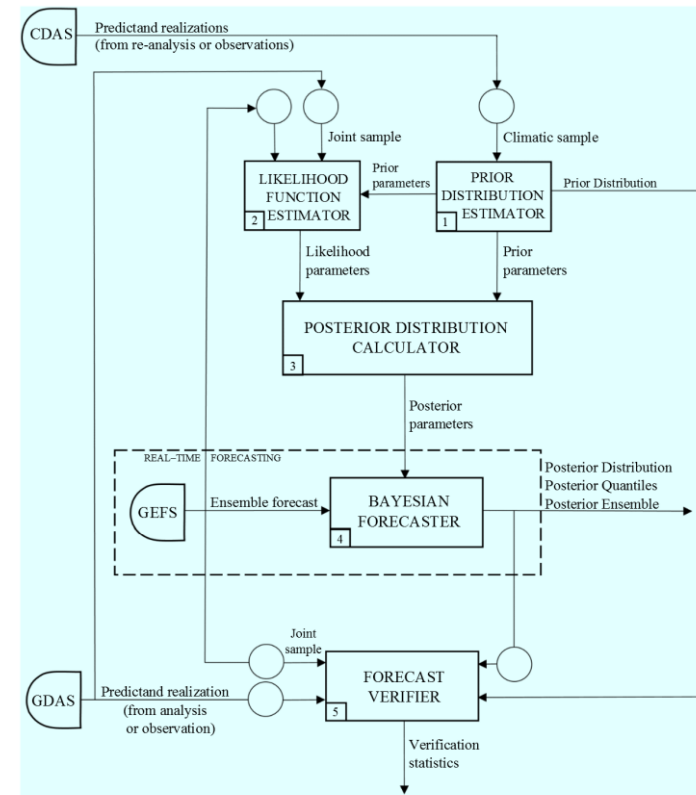   - "develop them for operations"

Figure 1. Structure of the BPE — its five components and its couplings with systems supplying data, illustrated for NCEP: Global Ensemble Forecast System (GEFS), Global Data Analysis System (GDAS), and Climate Data Analysis

Toth et al, GSD/ESRL

- Summary research project results

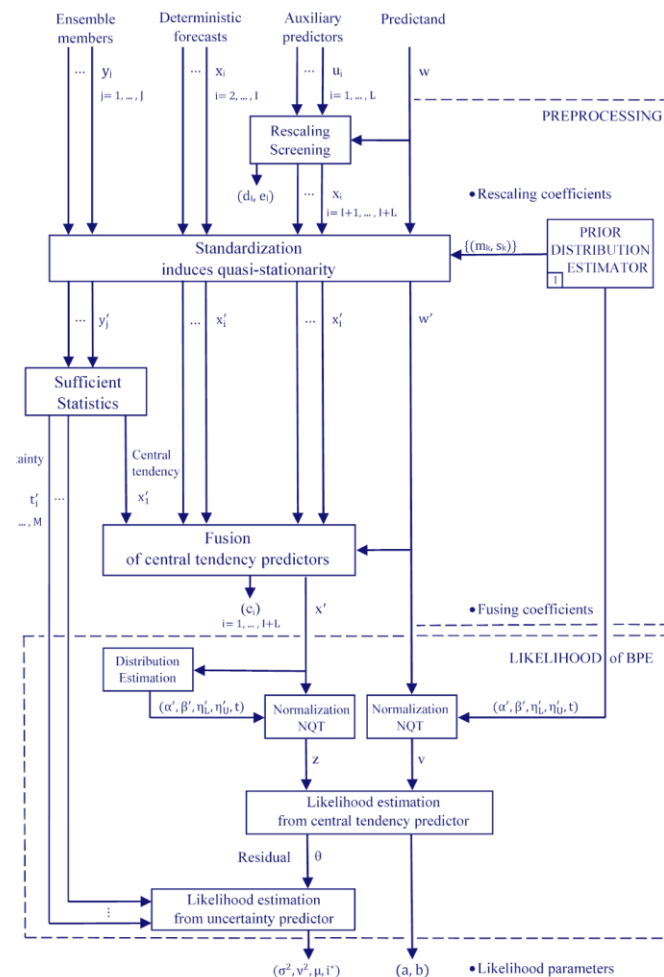- Potential transition to operations/future implementation

Figure 2. Structure of the likelihood function estimator of the BPE, version **SMM**, wh[...] processes and fuses **S**ingle ensemble forecast, **M**ultiple deterministic forecas[...] **M**ultiple auxiliary predictors.

Weather Information Statistical Post-processing System

Matt Peroutka, NOAA/NWS/STI

- Community-based software system, designed to statistically post-process NWP output throughout the weather enterprise

- Modern software techniques and self-describing data formats

- Widely-accepted standards

- Facilitate transition from research to operations

- No NGGPS funding (NWS base to date)
- Clear connection